On the LOCC Classification of Bipartite Density Matrices

Patrick Hayden[1], Barbara M. Terhal[2] and Armin Uhlmann[3]

[1]*Centre for Quantum Computation, Clarendon Laboratory, Parks Road, Oxford, OX1 3PU, UK;*
[2]*IBM Watson Research Center, P.O. Box 218, Yorktown Heights, NY 10598, US;*
[3]*Institut f. Theoretische Physik, Universität Leipzig, Augustusplatz 10/11, D-0409, Leipzig;*
Emails: `patrick.hayden@qubit.org`, `terhal@watson.ibm.com`, `armin.uhlmann@itp.uni-leipzig.de`
(February 1, 2008)

We provide a unifying framework for exact, probabilistic, and approximate conversions by local operations and classical communication (LOCC) between bipartite states. This framework allows us to formulate necessary and sufficient conditions for LOCC conversions from pure states to mixed states and it provides necessary conditions for LOCC conversions between mixed states. The central idea is the introduction of convex sets for exact, probabilistic, and approximate conversions, which are closed under LOCC operations and which are largely characterized by simple properties of pure states.

I. INTRODUCTION

Bipartite entanglement has been shown to be a valuable resource in quantum information theory, allowing two parties, Alice and Bob to teleport unknown quantum states [1], to remotely prepare quantum states [2,3] using classical communication only or to carry out other quantum information processing tasks. It is therefore natural that a theory of bipartite entanglement has been developed which attempts to quantify this resource. In this theory, a central role is played by the set of quantum operations which are constructed from local quantum operations by the parties Alice and Bob and classical communication between Alice and Bob, since these operations cannot enhance the quantum correlations in a bipartite state. Here we will call these quantum operations LOCC operations, for Local Operations and Classical Communication. The goal of the theory of bipartite entanglement is to obtain a classification of states with respect to the set of LOCC operations.

Since the notion of majorization [4] is to play a central role in what follows, let us recall some helpful notation. Let $\vec{\lambda}$ and $\vec{\mu}$ be two n-dimensional vectors with real coefficients in decreasing order, $\lambda^1 \geq \lambda^2 \geq \ldots \geq \lambda^n$, $\sum_{i=1}^{n} \lambda^i = 1$ and likewise for $\vec{\mu}$. We write

$$\vec{\lambda} \succ \vec{\mu} \Leftrightarrow \forall\, k = 1, \ldots, n \sum_{i=1}^{k} \lambda^i \geq \sum_{i=1}^{k} \mu^i. \tag{1}$$

When for all l we have

$$p \sum_{i=l}^{n} \lambda^i \leq \sum_{i=l}^{n} \mu^i, \tag{2}$$

we say that $p\vec{\lambda} \succ^w \vec{\mu}$, or in words, $\vec{\mu}$ is weakly supermajorized by $p\vec{\lambda}$. For a bipartite pure state $|\psi\rangle$ we write $\vec{\lambda}_\psi$ for the vector of Schmidt coefficients of $|\psi\rangle$ *in decreasing order*. (If $|\psi\rangle$ is unnormalized then $\vec{\lambda}_\psi$ is associated with the eigenvalues of $(\mathrm{Tr}_A|\psi\rangle\langle\psi|)/\langle\psi|\psi\rangle$, again in decreasing order.) In this paper we will call $|\langle\psi_1|\psi_2\rangle|$ and $F(\rho, \rho') = \mathrm{Tr}(\sqrt{\rho^{1/2}\rho'\rho^{1/2}})$ the square-root-fidelity between states $|\psi_1\rangle$ and $|\psi_2\rangle$ and ρ and ρ' respectively.

Let us begin our study by recapturing what is known for LOCC conversions between pure states. Let $|\psi_{\vec{\mu}}\rangle$ be a state that we start with, which is characterized by a Schmidt vector $\vec{\mu}$.

1. Nielsen [5]: A pure state $|\psi\rangle$ can be obtained by LOCC from the state $|\psi_{\vec{\mu}}\rangle$ *exactly* if and only if

$$\vec{\lambda}_\psi \succ \vec{\mu}. \tag{3}$$

2. Vidal [6]: A pure state $|\psi\rangle$ can be obtained by LOCC from the state $|\psi_{\vec{\mu}}\rangle$ *with probability* p if and only if

$$p\vec{\lambda}_\psi \succ^w \vec{\mu}. \tag{4}$$

The protocol that achieves the optimal probability can be viewed as a two-step protocol

$$|\psi_{\vec{\mu}}\rangle \rightarrow_{exact} |\xi\rangle \rightarrow_{\mathcal{M}_A} |\psi\rangle, \tag{5}$$

where *exact* and $\mathcal{M}_A$ denote an exact conversion and a unilateral measurement by Alice respectively. The state $|\xi\rangle$ is an intermediate state which can be determined by a procedure given in Ref. [6].

3. Vidal-Jonathan-Nielsen [7]: The optimal $f = |\langle\psi|\psi'\rangle|$ with which we approximate a state $|\psi\rangle$ by a state $|\psi'\rangle$ which we obtain by LOCC from $|\psi_{\vec{\mu}}\rangle$ is equal to $|\langle\psi|\xi\rangle|$ where $|\xi\rangle$ is the intermediate state in the optimal probabilistic conversion of $|\psi_{\vec{\mu}}\rangle$ into $|\psi\rangle$. (Note that in cases where the optimal probability is zero, the protocol still allows for the construction of the state $|\xi\rangle$.)

Our goal in this paper is to extend these results to the domain of mixed states. More specifically, we are interested in the 'formation' problem for bipartite mixed states: Given a pure state $|\psi_{\vec{\mu}}\rangle$, when can we obtain a density matrix ρ from $|\psi_{\vec{\mu}}\rangle$ by LOCC, *exactly* or *with probability p*? Similarly, what is the maximum fidelity we can achieve when *approximating* ρ? The problem of entanglement distillation [8] for a mixed states is not covered by this formalism. What we will find is that the 'formation' problem of mixed states can be completely translated to what is known for pure states, *i.e.* the 3 items listed above. Our results are expressed in four theorems that we present in this section and that we will prove later on in the paper.

Let us first define a restricted class of LOCC operations. We will call a quantum operation an LOCC_1^A operation when it is given by a local measurement/superoperator by Alice consisting of a set of commuting Kraus operators, followed by 1-way classical communication from Alice to Bob, and additional local unitary transformations by Alice and Bob, which may depend on Alice's measurement outcomes. Similarly, the class of LOCC_1^B maps is the set of LOCC_1^A maps with Alice and Bob interchanged.

For every pure state $|\psi_{\vec{\mu}}\rangle$ we define three kinds of sets, which we call $S_{\vec{\mu}}^+$, $S_{\vec{\mu},p}^+$ and $S_{\vec{\mu},f}^+$, where the first set relates to exact conversions, the second to probabilistic conversions and the last one to approximate conversions. These sets are defined in the following way

Definition 1 (Exact) *The set $S_{\vec{\mu}}$ is the convex closure of the set of pure states $|\psi\rangle$ such that*

$$\vec{\lambda}_\psi \succ \vec{\mu}. \tag{6}$$

The set $S_{\vec{\mu}}^+$ is defined as

$$S_{\vec{\mu}}^+ = \{\mathcal{L}(|\psi\rangle\langle\psi|) \,|\, \mathcal{L} \in \text{LOCC}_1^A, \ |\psi\rangle \in S_{\vec{\mu}}\}. \tag{7}$$

Note that the sets $S_{\vec{\mu}}$ are generalizations of the Schmidt number sets S_k which were introduced in Ref. [9]. A Schmidt number set S_k is a set $S_{\vec{\mu}}$ where $\vec{\mu}$ are the Schmidt coefficients of a maximally entangled state of Schmidt rank k.

Definition 2 (Probabilistic) *The set $S_{\vec{\mu},p}$ is the convex closure of the set of pure states $|\psi\rangle$ such that*

$$p\vec{\lambda}_\psi \succ^w \vec{\mu}. \tag{8}$$

The set $S_{\vec{\mu},p}^+$ is defined as

$$S_{\vec{\mu},p}^+ = \{\mathcal{L}(|\psi\rangle\langle\psi|) \,|\, \mathcal{L} \in \text{LOCC}_1^A, \ |\psi\rangle \in S_{\vec{\mu},p}\}. \tag{9}$$

In Section II C 2 we will discuss the relation between the sets $S_{\vec{\mu},p}^+$ and S_k.

Definition 3 (Approximate) *The set $S_{\vec{\mu},f}$ is the convex closure of the set of pure states $|\psi\rangle$ such that there exists a pure state $|\xi\rangle \in S_{\vec{\mu}}$ with $|\langle\xi|\psi\rangle| \geq f$. The set $S_{\vec{\mu},f}^+$ is defined as*

$$S_{\vec{\mu},f}^+ = \{\mathcal{L}(|\psi\rangle\langle\psi|) \,|\, \mathcal{L} \in \text{LOCC}_1^A, \ |\psi\rangle \in S_{\vec{\mu},f}\}. \tag{10}$$

It is not hard to see that the sets given in the Definitions 1-3 are all compact, convex sets. In Appendix A we will prove that unlike the Schmidt number sets, these sets do not relate directly to positive linear maps. Nonetheless, the characterization of these sets in terms of a pure state and an LOCC$_1^A$ operation is extremely simple. We are able to prove the following three theorems. The first is essentially a re-statement of a result by Jonathan and Plenio [10] in the language of $S_{\vec{\mu}}^+$ sets.

Theorem 1 *A density matrix $\rho \in S_{\vec{\mu}}^+$ if and only if there exists a decomposition $\{p_i, |\psi_i\rangle\}$ of ρ such that*

$$\sum_i p_i \vec{\lambda}_{\psi_i} \succ \vec{\mu}. \tag{11}$$

Furthermore, the LOCC conversion

$$|\psi_{\vec{\mu}}\rangle \rightarrow_{exact} \rho, \tag{12}$$

is possible if and only if $\rho \in S_{\vec{\mu}}^+$.

A straightforward corollary of this theorem is

Corollary 1 *Let $\rho \in S_{\vec{\mu}}^+$, then $E(\rho) \leq H(\vec{\mu})$ where $H(\vec{\mu})$ is the Shannon entropy of the vector $\vec{\mu}$ and E is the entanglement of formation of ρ,*

$$E(\rho) = \min_{\mathcal{E}=\{p_i, |\psi_i\rangle : \rho = \sum_i p_i |\psi_i\rangle\langle\psi_i|\}} \sum_i p_i E(|\psi_i\rangle\langle\psi_i|), \tag{13}$$

where $E(|\psi\rangle\langle\psi|)$ is the von Neumann entropy of the reduced density matrix of $|\psi\rangle$.

The second theorem, dealing with the case of probabilistic conversions from pure to mixed states, combines the results in Ref. [10] on exact conversions for mixed states with those of Vidal for probabilistic conversions between pure states [6].

Theorem 2 *A density matrix $\rho \in S_{\vec{\mu},p}^+$ if and only if there exists a decomposition $\{p_i, |\psi_i\rangle\}$ of ρ such that*

$$p \sum_i p_i \vec{\lambda}_{\psi_i} \succ^w \vec{\mu}. \tag{14}$$

Furthermore, the LOCC conversion

$$|\psi_{\vec{\mu}}\rangle \rightarrow \rho, \tag{15}$$

is possible with probability at least p if and only if $\rho \in S_{\vec{\mu},p}^+$.

For the last theorem, we need to introduce the notion of $(\vec{\mu}, f)$-approximability.

Definition 4 *A density matrix ρ is $(\vec{\mu}, f)$-approximable if and only if*

$$\sum_i |\langle\psi_i|\psi_i'\rangle| \geq f, \tag{16}$$

for some decomposition $\rho = \sum_i |\psi_i\rangle\langle\psi_i|$, and set of states $\{|\psi_i'\rangle\}$ such that $\mathrm{Tr}\sum_i |\psi_i'\rangle\langle\psi_i'| = 1$, $p_i = \langle\psi_i'|\psi_i'\rangle$, and

$$\sum_i p_i \vec{\lambda}_{\psi_i'} \succ \vec{\mu}. \tag{17}$$

Remarks We could have associated a density matrix $\rho' = \sum_i |\psi_i'\rangle\langle\psi_i'| \in S_{\vec{\mu}}^+$ with the set of approximating states $\{|\psi_i'\rangle\}$. It is not known whether our definition of $(\vec{\mu}, f)$-approximability is equivalent with the more natural definition in terms of the square root of the transition probability $F(\rho, \rho') = \mathrm{Tr}(\sqrt{\rho^{1/2}\rho'\rho^{1/2}})$; we could have called a state ρ $(\vec{\mu}, f)$-approximable iff there exists a $\rho' \in S_{\vec{\mu}}^+$ such that $F(\rho, \rho') \geq f$. Because of the strong concavity property of the square-root-fidelity [11],

$$F(\rho, \rho') \geq \sum_i |\langle \psi_i | \psi_i' \rangle| \tag{18}$$

whenever ρ is $(\vec{\mu}, f)$-approximable according to Definition 4, it follows that ρ is also approximable according to a definition which uses the square-root-fidelity. Furthermore, even though there exists a pair of decompositions which saturate Eq. (18), see Ref. [12], these optimal decompositions need not correspond to ones which obey the relation of Eq. (17). While we will further investigate the connection between F and $(\vec{\mu}, f)$-approximability in Section II D, it remains unclear whether this second definition would allow for Theorem 3.

The following theorem and its corollaries, when combined with the Vidal-Jonathan-Nielsen results [7], create a framework for answering the question of approximating a mixed state using LOCC operations, starting from a given pure state.

Theorem 3 *A density matrix $\rho \in S_{\vec{\mu},f}^+$ if and only if ρ is $(\vec{\mu}, f)$-approximable.*

It turns out that for two qubits a simplification takes place:

Theorem 4 *For two qubits $S_{\vec{\mu},x}^+ = S_{\vec{\mu},x}$ where x refers to either exact, probabilistic or approximate conversions. Let S be the set of all bipartite 2-qubit density matrices. We have*

$$0 < q \leq 2\mu_2 : \quad S_{(\mu_1,\mu_2),p=q} = S, \tag{19}$$

and

$$1 \geq q > 2\mu_2 : \quad S_{(\mu_1,\mu_2),p=q} = S_{(\mu_2/q, 1-\mu_2/q)}. \tag{20}$$

Thus for two qubits, the probabilistic sets are identical to each other and identical to the exact sets.

Theorems 1-3 also give rise to *necessary* conditions for LOCC conversions from a mixed state to either a pure or mixed state. From Theorem 1 we can conclude that

Corollary 2 *The LOCC conversion*

$$\rho_1 \rightarrow_{exact} \rho_2, \tag{21}$$

where ρ_1 is a mixed state is not possible if there exists a pure state $|\psi_{\vec{\mu}}\rangle$ such that $\rho_1 \in S_{\vec{\mu}}^+$, but $\rho_2 \notin S_{\vec{\mu}}^+$.

For probabilistic conversions, Theorem 2 leads immediately to

Corollary 3 *Suppose that $\rho_1 \in S_{\vec{\mu},p}^+$ and that, in particular, it is possible to obtain ρ_1 from $|\psi_{\vec{\mu}}\rangle$ with some probability $p_1 \geq p$. Now suppose $\rho_2 \notin S_{\vec{\mu},p}^+$. The maximum probability of success p_2 of obtaining ρ_2 from ρ_1 via LOCC operations satisfies $p_2 < p/p_1$.*

II. LOCC CONVERSIONS

A. Preliminaries

In proving Theorems 1 and 2 one observation turns out to be crucial.

Observation 1 When a density matrix ρ can be obtained by LOCC from a pure state $|\psi_{\vec{\mu}}\rangle$ with probability $p \in (0, 1]$ then there exists some decomposition $\{p_i, |\psi_i\rangle\}$ of ρ into pure states, *i.e.* $\rho = \sum_i p_i |\psi_i\rangle\langle\psi_i|$, such that we can obtain the ensemble of states $|\psi_i\rangle$ each with probability pp_i by LOCC from the state $|\psi_{\vec{\mu}}\rangle$.

It is essential that the LOCC protocol starts with a pure state. Consider a unilateral operation by, say, Alice, on a pure state $|\psi\rangle$:

$$|\psi\rangle \rightarrow \{p_i, \mathcal{S}_i(|\psi\rangle\langle\psi|)\}, \tag{22}$$

where $\mathcal{S}_i$ is some CP map, characterized by Kraus operators $\{A_k^i\}$. Instead of performing $\mathcal{S}_i$, she can perform the more 'fine-grained' measurement

$$|\psi\rangle \rightarrow \left\{ p_i q_{ij}, \frac{|\psi_{ij}\rangle\langle\psi_{ij}|}{q_{ij}} \right\}, \tag{23}$$

where $|\psi_{ij}\rangle = A^i_j \otimes \mathbf{1}|\psi\rangle$ and $q_{ij} = \mathrm{Tr}\, A^{i\,\dagger}_j A^i_j \otimes \mathbf{1}|\psi\rangle\langle\psi|$, which gives her an ensemble of pure states. The extra bits of information that she obtains will be communicated to Bob in the next round. By replacing every 'mixing' operator in the LOCC protocol by a fine-grained measurement projecting on pure states, Alice and Bob obtain a new protocol which produces an ensemble of pure states as output, and the ensemble forms a decomposition of the mixed state that was obtained in the original LOCC protocol.

B. Exact

1. Proof of Theorem 1

The proof of this Theorem follows quite straightforwardly from Observation 1 and the results of Jonathan and Plenio [10]. Theorem 1 of Ref. [10] states that we can obtain the ensemble $\{p_i, |\psi_i\rangle\}$ ($\sum_i p_i = 1$) by LOCC from a state $|\psi_{\vec{\mu}}\rangle$ if and only if

$$\sum_i p_i \vec{\lambda}_{\psi_i} \succ \vec{\mu}. \tag{24}$$

Now, if we can obtain the ensemble $\{p_i, |\psi_i\rangle\}$ by LOCC, then we can obtain $\rho = \sum_i p_i |\psi_i\rangle\langle\psi_i|$. On the other hand, if we can obtain ρ by LOCC from the state $|\psi_{\vec{\mu}}\rangle$, then we can obtain *some* decomposition of ρ by Observation 1, and thus for this ensemble Eq. (24) should hold.

The equivalence between Eq. (24) for some ensemble $\{p_i, |\psi_i\rangle\}$ of ρ and $\rho \in S^+_{\vec{\mu}}$ follows from inspecting the Jonathan-Plenio [10] protocol for converting $|\psi_{\vec{\mu}}\rangle$ to the ensemble $\{p_i, |\psi_i\rangle\}$. It is a two step protocol:

$$|\psi_{\vec{\mu}}\rangle \rightarrow_{exact} |\bar{\psi}\rangle \rightarrow_{\mathrm{LOCC}^A_1} \rho, \tag{25}$$

where $|\bar{\psi}\rangle \in S_{\vec{\mu}}$ and its Schmidt vector $\vec{\lambda}_{\bar{\psi}}$ is given by $\lambda^i_{\bar{\psi}} = \sum_j p_j \lambda^i_{\psi_j}$. Thus, every ρ for which there exists an ensemble obeying Eq. (24), can be written as $\rho = \mathcal{L}(|\psi\rangle\langle\psi|)$ where $\mathcal{L}$ is some LOCC^A_1 map and $|\psi\rangle \in S_{\vec{\mu}}$. These observations prove Theorem 1. $\square$

2. Consequences

Note that since the majorization conditions treat Alice and Bob on an equal basis, we can alternatively write

$$S^+_{\vec{\mu}} = \{ \mathcal{L}(|\psi\rangle\langle\psi|) \,|\, \mathcal{L} \in \mathrm{LOCC}^B_1, \ |\psi\rangle \in S_{\vec{\mu}} \}. \tag{26}$$

One might ask whether the extra complication incurred by defining the sets $S^+_{\vec{\mu}}$ is necessary. While we will find later on that in the case of pairs of qubits the sets $S^+_{\vec{\mu}}$ are equal to the simpler sets $S_{\vec{\mu}}$, the following example illustrates that for pairs of n-level systems, with $n \geq 3$, that the sets generally do not coincide.

Example 1 *Let $|\psi_0\rangle = |00\rangle$ and $|\psi_{12}\rangle = \frac{1}{\sqrt{2}}(|11\rangle + |22\rangle)$ then consider the density operator $\rho = (1-\epsilon)|\psi_0\rangle\langle\psi_0| + \epsilon|\psi_{12}\rangle\langle\psi_{12}|$. Let $\vec{\mu} = (1 - \epsilon/2, \epsilon/2, 0)$. Then for $\epsilon \in (0, 1)$, $\rho \in S^+_{\vec{\mu}}$ but $\rho \notin S_{\vec{\mu}}$.*

First, notice that $\vec{\lambda}_{\psi_0} = (1, 0, 0)$ while $\vec{\lambda}_{\psi_{12}} = (\frac{1}{2}, \frac{1}{2}, 0)$. Thus, $(1-\epsilon)\vec{\lambda}_{\psi_0} + \epsilon\vec{\lambda}_{\psi_{12}} = \vec{\mu}$ and we have that $\rho \in S^+_{\vec{\mu}}$.

Now we need to demonstrate that for any decomposition $\rho = \sum_i p_i |\phi_i\rangle\langle\phi_i|$ there always exists an i such that $\vec{\lambda}_{\phi_i} \not\succ \vec{\mu}$. Let

$$P_{12} = (|1\rangle\langle1| + |2\rangle\langle2|) \otimes (|1\rangle\langle1| + |2\rangle\langle2|), \tag{27}$$

and $P_0 = |00\rangle\langle00|$. Then $P_{12}\rho P_{12} = \epsilon|\psi_{12}\rangle\langle\psi_{12}|$ from which it follows that $P_{12}|\phi_i\rangle\langle\phi_i|P_{12} \propto |\psi_{12}\rangle\langle\psi_{12}|$ and similarly $P_0|\phi_i\rangle\langle\phi_i|P_0 \propto |00\rangle\langle00|$. Together this implies that

$$|\phi_i\rangle = (P_0 + P_{12})|\phi_i\rangle = \alpha_i|00\rangle + \frac{\beta_i}{\sqrt{2}}|11\rangle + \frac{\beta_i}{\sqrt{2}}|22\rangle, \tag{28}$$

for some complex amplitudes α_i, β_i. Letting $\alpha^*_i \alpha_i = r_i$ and $\beta^*_i \beta_i = s_i$, we write $\vec{\lambda}_{\phi_i} = (r_i, s_i/2, s_i/2)$, assuming $r_i > s_i/2$. If both r_i and s_i are nonzero, then $\vec{\lambda}_{\phi_i} \not\succ \vec{\mu}$. On the other hand, any decomposition of ρ into pure states each with either $r_i = 0$ or $s_i = 0$ will necessarily contain terms for which $|\phi_i\rangle = |\psi_{12}\rangle$. Thus, $\rho \notin S_{\vec{\mu}}$. $\square$

C. Probabilistic

1. Proof of Theorem 2

First, we prove that if a state ρ has a decomposition $\rho = \sum_j p_j |\psi_j\rangle\langle\psi_j|$ satisfying the condition

$$p \sum_j p_j \vec{\lambda}_{\psi_j} \succ^w \vec{\mu}, \tag{29}$$

then $\rho \in S^+_{\vec{\mu},p}$. Let us define the state $|\bar{\psi}\rangle$ with Schmidt coefficients $\lambda^i = \sum_j p_j \lambda^i_{\psi_j}$. Eq. (29) directly implies that $|\bar{\psi}\rangle \in S_{\vec{\mu},p}$ and thus by Vidal's result on probabilistic pure state conversions, Eq. (4), we can obtain $|\bar{\psi}\rangle$ with probability p by LOCC from $|\psi_{\vec{\mu}}\rangle$. Then an LOCC^A_1 map as in the last step of Eq. (25) results in ρ. This shows that that $\rho \in S^+_{\vec{\mu},p}$ (and also that we can obtain ρ by LOCC with probability p).

On the other hand, let $\rho \in S^+_{\vec{\mu},p}$, i.e. $\rho = \mathcal{L}(|\psi\rangle\langle\psi|)$ with $\mathcal{L} \in \mathrm{LOCC}^1_A$ and $|\psi\rangle \in S_{\vec{\mu},p}$. This implies that we can obtain ρ with probability at least p by LOCC from $|\psi_{\vec{\mu}}\rangle$. By Theorem 1 we have that there exists a decomposition $\{p_i, |\psi_i\rangle\}$ of ρ such that

$$\sum_i p_i \vec{\lambda}_{\psi_i} \succ \vec{\lambda}_\psi, \tag{30}$$

and by Eq. (4), we get

$$p \sum_i p_i \vec{\lambda}_{\psi_i} \succ p\vec{\lambda}_\psi \succ^w \vec{\mu}, \tag{31}$$

which is the desired condition. $\square$

As before, this result allows us to write down an alternative characterization of the sets $S^+_{\vec{\mu},p}$, namely

$$S^+_{\vec{\mu},p} = \{\mathcal{L}(|\psi\rangle\langle\psi|) \,|\, \mathcal{L} \in \mathrm{LOCC}^B_1, \, |\psi\rangle \in S_{\vec{\mu},p}\}. \tag{32}$$

2. Relation with Schmidt number sets

What is the relation of the sets $S^+_{\vec{\mu},p}$ with the Schmidt number sets S_k? For every vector $\vec{\mu}$, there is some probability p by which $|\psi_{\vec{\mu}}\rangle$ can be mapped onto $|\Psi_k\rangle$ where k is the Schmidt rank of $\vec{\mu}$. Therefore $S_k \subseteq S^+_{\vec{\mu},p_{max}}$ where p_{max} is given by the maximum probability p such that $p\vec{\lambda}_{\Psi_k} \succ^w \vec{\mu}$. Note also that $S^+_{\vec{\mu},p_{max}} \subseteq S^+_{\vec{\mu},q}$ where $q \leq p_{max}$. On the other hand, we have $S^+_{\vec{\mu},p} \subseteq S_k$ for all $p > 0$. This follows from the fact that (1) $|\psi_{\vec{\mu}}\rangle$ has Schmidt rank k and (2) any state $|\psi_\mu\rangle$ can be reached by LOCC from $|\Psi_k\rangle$ with probability 1, due to the condition in Eq. (3), and (3) from $|\psi_{\vec{\mu}}\rangle$ all states in $S^+_{\vec{\mu},p}$ can be obtained by LOCC with probability at least p.

Therefore

$$S_k = S^+_{\vec{\mu},q}, \; \forall \, 0 < q \leq p_{max}, \tag{33}$$

that is, for small probability up to some maximum, the sets $S^+_{\vec{\mu},q}$ collapse onto the set S_k and onto each other.

D. Approximate conversions

The following definition will be useful in what follows.

Definition 5 (Optimal $\vec{\mu}$-approximation) *A density matrix $\rho' = \sum_i |\psi'_i\rangle\langle\psi'_i|$, where $p_i = \langle\psi'_i|\psi'_i\rangle$ and*

$$\sum_i p_i \vec{\lambda}_{\psi'_i} \succ \vec{\mu}, \tag{34}$$

is an optimal $\vec{\mu}$-approximation *to $\rho = \sum_i |\psi_i\rangle\langle\psi_i|$ when*

$$\sum_i |\langle\psi_i|\psi'_i\rangle| \tag{35}$$

is maximal with respect to all decompositions of ρ and sets of states $|\psi'_i\rangle$ for which condition Eq. (34) holds.

Let ρ be a $(\vec{\mu}, f)$-approximable state and ρ' be an optimal $\vec{\mu}$-approximation to ρ. Thus there exist decompositions $\rho = \sum_i |\psi_i\rangle\langle\psi_i|$ and $\rho' = \sum_i |\psi_i'\rangle\langle\psi_i'|$ such that

$$\sum_i |\langle\psi_i|\psi_i'\rangle| \geq f. \tag{36}$$

Since $\rho' \in S_{\vec{\mu}}^+$, we can find unitary operators U_i' and V_i', non-negative diagonal Kraus operators $A_i = \text{diag}(\sqrt{\lambda_{i,1}}, \ldots, \sqrt{\lambda_{i,n}})$ satisfying $\sum_i A_i^2 = \mathbf{1}$ and a state $|\psi_{\rho'}\rangle \in S_{\vec{\mu}}$ such that $|\psi_i'\rangle = (U_i' \otimes V_i' A_i)|\psi_{\rho'}\rangle$. Similarly, we can find unitary U_i and V_i, operators $B_i = \text{diag}(\sqrt{\mu_{i,1}}, \ldots, \sqrt{\mu_{i,n}})$, $\sum_i B_i^2 = \mathbf{1}$ and a state $|\psi_\rho\rangle$ such that $|\psi_i\rangle = (U_i \otimes V_i B_i)|\psi_\rho\rangle$. Furthermore, the Plenio-Jonathan protocol requires that $|\psi_\rho\rangle$ and $|\psi_{\rho'}\rangle$ be in Schmidt form in the bases that diagonalize their respective Kraus operators A_i and B_i, which in both cases is the standard basis. Therefore, $|\psi_{\rho'}\rangle = \sum_k \sqrt{\alpha_k}|k,k\rangle$ and $|\psi_\rho\rangle = \sum_k \sqrt{\beta_k}|k,k\rangle$.

By Lemma 1 of Ref. [7], the absolute value of the inner product between two states $|\psi\rangle$ and $|\psi'\rangle$ with given Schmidt coefficients is maximized when the states are chosen to have the same Schmidt basis. We can therefore assume that $U_i' = U_i$ and $V_i' = V_i$ since ρ' is an optimal $\vec{\mu}$-approximation to ρ. We can then calculate

$$\sum_i |\langle\psi_i|\psi_i'\rangle| = \sum_{i,k} \sqrt{\alpha_k \lambda_{ik}}\sqrt{\beta_k \lambda_{ik}}. \tag{37}$$

However, since $\sum_i \lambda_{i,k} = \sum_i \mu_{i,k} = 1$, we know that $\sum_i \sqrt{\lambda_{i,k}\mu_{i,k}} \leq 1$. Therefore, again using the optimality of ρ', we can conclude that $\mu_{i,k} = \lambda_{i,k}$. It follows that

$$|\langle\psi_\rho|\psi_{\rho'}\rangle| = \sum_i |\langle\psi_i|\psi_i'\rangle| \geq f. \tag{38}$$

Therefore $|\psi_\rho\rangle \in S_{\vec{\mu},f}$. We have shown that when ρ is $(\vec{\mu}, f)$-approximable, we can write $\rho = \mathcal{L}(|\psi_\rho\rangle\langle\psi_\rho|)$ where $|\psi_\rho\rangle \in S_{\vec{\mu},f}$ and $\mathcal{L}$ is an LOCC$_1^A$ map.

On the other hand, let $\rho \in S_{\vec{\mu},f}^+$; we can write $\rho = \mathcal{L}(|\psi\rangle\langle\psi|)$ where $|\psi\rangle \in S_{\vec{\mu},f}$. Consider an optimal $\vec{\mu}$-approximation $|\psi'\rangle \in S_{\vec{\mu}}$ to $|\psi\rangle$. As an approximation to ρ we take $\rho' = \mathcal{L}(|\psi'\rangle\langle\psi'|)$. We write $\rho = \sum_i |\psi_i\rangle\langle\psi_i|$ and $\rho' = \sum_i |\psi_i'\rangle\langle\psi_i'|$ where $(U_i \otimes A_i)|\psi\rangle = |\psi_i\rangle$ and similarly for $|\psi_i'\rangle$, where A_i are the Kraus operators of a local superoperator implemented by Alice. Then we have

$$\sum_i |\langle\psi_i|\psi_i'\rangle| = \sum_i |\langle\psi|A_i^\dagger A_i \otimes \mathbf{1}|\psi'\rangle| \geq |\langle\psi|\psi'\rangle| \geq f, \tag{39}$$

where we used $\sum_i |x_i| \geq |\sum_i x_i|$. Therefore ρ is $(\vec{\mu}, f)$-approximable. $\square$

2. Connection to square-root fidelity

We will now investigate in more detail the relationship between square-root fidelity and $(\vec{\mu}, f)$-approximability. For convenience, we begin with

Definition 6 ($\vec{\mu}$-fidelity) *Let ρ and σ be density operators. The $\vec{\mu}$-fidelity is defined to be $f_{\vec{\mu}}(\rho, \sigma) = \max \sum_i |\langle\phi_i|\psi_i\rangle|$, where $\rho = \sum_i |\phi_i\rangle\langle\phi_i|$ and $\sigma = \sum_i |\psi_i\rangle\langle\psi_i|$ with $\sum_i p_i \vec{\lambda}_{\psi_i} \succ \vec{\mu}$ for $p_i = \langle\psi_i|\psi_i\rangle$. If no such decomposition exists then set $f_{\vec{\mu}}(\rho, \sigma) = 0$.*

Furthermore, define the following two optimal fidelities:

$$F_{max}(\rho, \vec{\mu}) = \max\{F(\rho, \sigma) : \sigma \in S_{\vec{\mu}}^+\} \tag{40}$$

$$f_{max}(\rho, \vec{\mu}) = \max\{f_{\vec{\mu}}(\rho, \sigma) : \sigma \in S_{\vec{\mu}}^+\}. \tag{41}$$

In this section, we will develop some properties of F_{max} and f_{max} that might shed some light on the relationship between them. We can concisely state our results on approximation by noting that $F_{max}(\rho, \vec{\mu}) \geq f_{max}(\rho, \vec{\mu})$ and $\rho \in S_{\vec{\mu},f}^+$ if and only if $f_{max}(\rho, \vec{\mu}) \geq f$. However, since F_{max} is a more natural definition of the optimal fidelity of

approximation, we would ideally like to re-phrase the second result in terms of F_{max}. We will not be able to do so but can offer some suggestive partial results.

To begin, both functions are easily seen to be consistent with the partial ordering on Schmidt vectors: $\vec{\lambda} \succ \vec{\mu}$ implies that $F_{max}(\rho, \vec{\lambda}) \leq F_{max}(\rho, \vec{\mu})$, and similarly for f_{max}. Furthermore, notice that for a pure state $|\phi\rangle$ and Schmidt vector $\vec{\mu}$,

$$f_{max}(|\phi\rangle\langle\phi|, \vec{\mu}) = F_{max}(|\phi\rangle\langle\phi|, \vec{\mu}). \tag{42}$$

One way to see this is to recall the result from Ref. [7] that the best approximation with respect to F to a given pure state $|\phi\rangle$ by a mixed state $\sigma \in S_{\vec{\mu}}^{+}$ can always itself be taken to be given by a pure state $|\psi\rangle \in S_{\vec{\mu}}^{+}$. The result then follows because $f_{\vec{\mu}}(|\phi\rangle, |\psi\rangle)$ and $F(|\phi\rangle, |\psi\rangle)$ agree if $\lambda_{|\psi\rangle} \succ \vec{\mu}$.

The functions are similar in other ways as well, both them obeying a strong form of joint concavity. Let us begin with F_{max}. Suppose that $F_{max}(\rho_j, \vec{\mu}_j) = F(\rho_j, \sigma_j)$ for some $\sigma_j \in S_{\vec{\mu}_j}^{+}$. If we set $\vec{\mu} = \sum_j q_j \vec{\mu}_j$, then we find $\sum_j q_j \sigma_j \in S_{\vec{\mu}}^{+}$ so that

$$F_{max}\left(\sum_j p_j \rho_j, \sum_k q_k \vec{\mu}_k\right) \geq F\left(\sum_j p_j \rho_j, \sum_k q_k \sigma_k\right) \tag{43}$$

$$\geq \sum_j \sqrt{p_j q_j} F(\rho_j, \sigma_j) \tag{44}$$

$$= \sum_j \sqrt{p_j q_j} F_{max}(\rho_j, \vec{\mu}_j). \tag{45}$$

Let us now move on to the joint concavity of f_{max}, for which the proof is similar. Suppose that $f_{max}(\rho_j, \vec{\mu}_j) = f_{\vec{\mu}_j}(\rho_j, \sigma_j)$, where $\sigma_j \in S_{\vec{\mu}_j}^{+}$. Likewise, assume that the optimal decomposition is given by $f_{\vec{\mu}_j}(\rho_j, \sigma_j) = \sum_i |\langle\phi_{ij}|\psi_{ij}\rangle|$, where $\sum_i r_{ij} \vec{\lambda}_{\psi_{ij}} \succ \vec{\mu}_j$ for $r_{ij} = \langle\psi_{ij}|\psi_{ij}\rangle$. Then, since $\sum_{ij} q_j r_{ij} \vec{\lambda}_{\psi_{ij}} \succ \sum_j q_j \vec{\mu}_j$, we find

$$f_{max}\left(\sum_j p_j \rho_j, \sum_k q_k \vec{\mu}_k\right) \geq \sum_{ij} \sqrt{p_j q_j} |\langle\phi_{ij}|\psi_{ij}\rangle| = \sum_j \sqrt{p_j q_j} f_{\vec{\mu}_j}(\rho_j, \sigma_j) = \sum_j \sqrt{p_j q_j} f_{max}(\rho_j, \mu_j), \tag{46}$$

which is the desired inequality.

Summarizing then, for a function g equal to F_{max} or f_{max}, we have

1. (Consistency with partial order) $\vec{\lambda} \succ \vec{\mu}$ implies $g(\rho, \vec{\lambda}) \leq g(\rho, \vec{\mu})$.

2. (Agreement on pure states) For a pure state $|\phi\rangle$ and Schmidt vector $\vec{\mu}$, $g(|\phi\rangle\langle\phi|, \vec{\mu})$ is equal to the maximum over $|\psi\rangle \in S_{\vec{\mu}}$ of $|\langle\phi|\psi\rangle|$.

3. (Joint concavity) If $\rho = \sum_i p_i |\phi_i\rangle\langle\phi_i|$ and $\vec{\mu} = \sum_j q_j \vec{\mu}_j$ then $g(\rho, \vec{\mu}) \geq \sum_i \sqrt{p_i q_i} g(\rho_i, \vec{\mu}_i)$.

There is a difference, however. For f_{max} we also know that given ρ and $\vec{\mu}$, we can find a decomposition $\rho = \sum_i p_i |\phi_i\rangle\langle\phi_i|$ and an ensemble $\{q_i, |\psi_i\rangle\}$ with $\sum_i q_i \vec{\lambda}_{\psi_i} \succ \vec{\mu}$ such that $f_{max}(\rho, \vec{\mu}) = \sum_i \sqrt{p_i q_i} |\langle\phi_i|\psi_i\rangle|$. Since $|\langle\phi_i|\psi_i\rangle| \leq f_{max}(|\phi_i\rangle\langle\phi_i|, \vec{\lambda}_{\psi_i})$ by definition, summing inequalities gives

$$f_{max}(\rho, \vec{\mu}) \leq \sum_i \sqrt{p_i q_i} f_{max}(\rho_i, \vec{\lambda}_{\psi_i}). \tag{47}$$

Combining with joint concavity then implies that

$$f_{max}(\rho, \vec{\mu}) = \sum_i \sqrt{p_i q_i} f_{max}(|\phi_i\rangle\langle\phi_i|, \vec{\lambda}_{\psi_i}). \tag{48}$$

Thus, f_{max} is the smallest function consistent with the three properties listed above.

This is quite a curious situation. The results of Ref. [12] immediately imply that the square-root fidelity $F(\rho, \sigma)$ (note that the arguments of F are different than those of F_{max}) is the smallest function g satisfying the following different list of properties:

1. (Agreement on pure states) For normalized $|\phi\rangle$ and $|\psi\rangle$, $g(|\phi\rangle\langle\phi|, |\psi\rangle\langle\psi|) = |\langle\phi|\psi\rangle|$.

2. (Joint concavity) If $\rho = \sum_i p_i \rho_i$ and $\sigma = \sum_j q_j \sigma_j$ then $g(\rho, \sigma) \geq \sum_i \sqrt{p_i q_i} g(\rho_i, \sigma_i)$.

Thus F arises as the smallest function satisfying a set of conditions which appear very similar to those conditions satisfied by F_{max} and f_{max}. It seems plausible, therefore, that F_{max}, which is defined in terms of F, might itself be the smallest function consistent with properties 1-3 of F_{max} and f_{max}. If that is the case, then $f_{max} = F_{max}$.

8

Theorem 3 does not tell us how to determine the maximum $\vec{\mu}$-fidelity with which we can approximate a density matrix ρ, given a starting state $|\psi_{\vec{\mu}}\rangle$. To answer this question will generally require a further optimization over decompositions of ρ. Let us consider an example.

Example 2 *Let $|\Psi_m\rangle = 1/\sqrt{m}\sum_{i=1}^m |i_A\rangle|i_B\rangle$ be a maximally entangled bipartite state of Schmidt rank m and let $\vec{\lambda}_m$ be the Schmidt vector of $|\Psi_m\rangle$. The best approximation to a state ρ achievable starting from $|\Psi_m\rangle$ using LOCC operations will have an optimal approximation with*

$$f_{max}(\rho, \vec{\lambda}_m) = (1 - E_{m+1}(\rho))^{1/2}, \tag{49}$$

where $E_l(\rho)$ is one of Vidal's entanglement monotones [6], defined as

$$E_l(\rho) = \min_{\mathcal{E}} \sum_i p_i E_l(|\psi_i\rangle\langle\psi_i|), \tag{50}$$

in which $E_l(|\psi_i\rangle\langle\psi_i|) = \sum_{k=l}^n \lambda_{\psi_i}^k$ and $\mathcal{E} = \{p_i, |\psi_i\rangle\}$ is an ensemble realizing ρ.

In Ref. [7], it is demonstrated that if $|\phi\rangle = \sum_{i=1}^n \sqrt{\beta_i}|i_A\rangle|i_B\rangle$, with $\beta_i \geq \beta_{i+1}$ then the best approximation to $|\phi\rangle$ achievable using LOCC operations and starting from the state $|\Psi_m\rangle$ has square-root-fidelity $f_{max} = (\sum_{i=1}^m \beta_i)^{1/2}$. Theorem 3 shows that the optimal approximation for mixed states is that derived from some Jonathan-Plenio precursor state of the form $\sum_i \sqrt{\sum_j p_j \lambda_{\psi_j}^i}|i_A\rangle|i_B\rangle$, where $\rho = \sum_j p_j|\psi_j\rangle\langle\psi_j|$. Therefore, the best approximation to the mixed state ρ has fidelity

$$f_{max}(\rho, \vec{\lambda}_m) = \max\left(\sum_j p_j \sum_{i=1}^m \lambda_{\psi_j}^i\right)^{1/2} = \left(1 - \min\sum_j p_j \sum_{i=m+1}^n \lambda_{\psi_j}^i\right)^{1/2}, \tag{51}$$

where the optimizations are to be performed over all decompositions of ρ. The last line of the above equation, however, is simply $(1 - E_{m+1}(\rho))^{1/2}$. □

E. Two Qubits: Proof of Theorem 4

1. Exact

First we relate the set $S_{\vec{\mu}}$ to the entanglement of formation:

Proposition 1 *The entanglement of formation E of a density matrix in $\mathbf{C}_2 \otimes \mathbf{C}_2$ is*

$$E(\rho) = \min_{\vec{\mu}\,|\,\rho\in S_{\vec{\mu}}} H_2(\vec{\mu}), \tag{52}$$

where $H_2(.)$ is the binary entropy function, i.e. $H_2(x) = -x\log x - (1-x)\log(1-x)$.

Proof The Wootters formula for the entanglement of formation [13] tells us that in the optimal decomposition $\{p_i, |\psi_i\rangle\}$ of ρ, every state $|\psi_i\rangle$ has an equal amount of entanglement. Therefore if ρ has entanglement $E(\rho)$, the density matrix ρ must be contained in the set $S_{\vec{\mu}}$ where $H_2(\vec{\mu}) = E(\rho)$. Conversely, if $\rho \in S_{\vec{\mu}}$ then $E(\rho) \leq H_2(\vec{\mu})$.□

From this Proposition it follows that $S_{\vec{\mu}}$ must be closed under LOCC and therefore $S_{\vec{\mu}} = S_{\vec{\mu}}^+$. Furthermore, for two qubits we know that the partial order on the pure states induced by majorization is a total order, which is characterized by a single parameter, the smallest eigenvalue μ_2 of the reduced density matrix of the pure state. Therefore we can characterize the vector $\vec{\mu}$ with this single parameter μ_2 and it follows that $S_{\mu_2} \subset S_{\mu_2'}$ when $\mu_2 < \mu_2'$.

As a corollary of this result we obtain that we can convert $|\psi_{\vec{\mu}}\rangle$ to ρ by LOCC if and only if $\rho \in S_{\vec{\mu}}$, which means that $E(\rho) \leq H(\vec{\mu})$. Therefore the minimal entanglement costs for preparing a single copy of ρ exactly is $E(\rho)$. This result has been independently found by Vidal [15].

2. Probabilistic

Let us now consider the probabilistic sets and prove that for 2 qubits $S^+_{\vec{\mu},p} = S_{\vec{\mu},p}$. First we note that the relation with the Schmidt number sets gives us the following:

$$S = S_2 = S_{(\mu_1,\mu_2),p=q}, \tag{53}$$

when $q \leq p_{max} = 2\mu_2$. Here S_2 is the Schmidt number 2 set, which is identical to the set of all bipartite two-qubit density matrices, S. Now let $q > 2\mu_2$. A density matrix $\rho \in S^+_{\vec{\mu},q}$ if and only if there exists a decomposition $\{p_i, |\psi_i\rangle\}$ of ρ such that

$$q \sum_i p_i \lambda^2_{\psi_i} \leq \mu_2, \tag{54}$$

where $\lambda^2_{\psi_i}$ is the smallest Schmidt coefficients of $|\psi_i\rangle$. Since $\mu_2/q < 1/2$, this condition is identical to the requirement that $\rho \in S^+_{\vec{\mu}'=(\mu_2/q,1-\mu_2/q)}$. Since $S^+_{\vec{\mu}} = S_{\vec{\mu}}$ for two qubits, this condition is again identical to requirement that there exists a decomposition $\{p_i, |\psi_i\rangle\}$ of ρ such that for all i

$$q\lambda^2_{\psi_i} \leq \mu_2, \tag{55}$$

which implies that $\rho \in S_{\vec{\mu},q}$. So we have shown that for $q > 2\mu_2$, the exact sets $S_{(\mu_2/q,1-\mu_2/q)}$ and the probabilistic sets $S_{(\mu_1,\mu_2),q}$ are, in fact, identical.

3. Approximate

Finally, let us move on to the approximation sets. We'll begin by considering pure states. Suppose that $|\psi\rangle \notin S_{\vec{\mu}}$ and let $|\psi'\rangle \in S_{\vec{\mu}}$ such that $|\langle\psi|\psi'\rangle| = f_{max}$ is optimal. The condition that $|\psi\rangle \in S_{\vec{\mu},f}$ is that $f_{max} \geq f$. Suppose now that $|\psi\rangle$ has Schmidt coefficients $(\alpha, 1-\alpha)$, where $\alpha \geq 1/2$. It is clear that f_{max} is a strictly increasing function of α on the interval $[1/2, \mu_1]$, from which it follows that $f_{max} \geq f$, or, equivalently, that $|\psi\rangle \in S_{\vec{\mu},f}$, if and only if $\alpha \geq \mu'_1$ for some $\mu'_1 \in [1/2, \mu_1]$. If we set $\mu'_2 = 1 - \mu'_1$, it follows immediately that $S_{\vec{\mu},f} = S_{\vec{\mu}'}$ since the extreme points of these convex sets are pure states. It then follows from the definitions that $S^+_{\vec{\mu},f} = S^+_{\vec{\mu}'}$. We saw above, however, that $S^+_{\vec{\mu}'} = S_{\vec{\mu}'}$. Thus, $S^+_{\vec{\mu},f} = S_{\vec{\mu},f}$.

III. CONCLUSION

We have provided a unifying framework for exact, probabilistic, and approximate LOCC conversions from pure states to mixed states. In each case we have found criteria defining exactly when these conversions are possible, with the caveat that the criteria are always expressed in terms of the existence of 'optimal' decompositions of the target mixed state having some easily verified property. This work does not address the question of how to find these optimal decompositions of a density matrix ρ for exact, probabilistic or approximate conversions of an initial state $|\psi_{\vec{\mu}}\rangle$ into the state ρ. This problem could be as hard a determining optimal decompositions for the entanglement of formation, but may well be simpler. In addition to resolving that question, with our framework established, a host of open questions present themselves. The most pressing is the question of determining whether the function $f_{max}(\rho, \vec{\mu})$ we have been studying here is or is not equal to the more natural $F_{max}(\rho, \vec{\mu})$. Another question of interest is the relation between the sets $S^+_{\vec{\mu}}$ and the optimal decomposition of ρ with respect to its entanglement of formation.

IV. ACKNOWLEDGMENTS

APPENDIX A: POSITIVE LINEAR MAPS

We may ask whether, say, the sets $S_{\vec{\mu}}$ or $S_{\vec{\mu}}^+$ are characterized by positive linear maps. Let us define a $\vec{\mu}$-positive linear map in the following way, as a generalization of k-positivity. Here $\vec{\mu}$ is a n-dimensional Schmidt vector. We would propose the following definition. The linear Hermiticity-preserving map $\mathcal{L}$ is positive with respect to $|\psi_{\vec{\mu}}\rangle$, or $\vec{\mu}$-positive if $(\mathbf{1} \otimes \mathcal{L})(\rho) \geq 0$ for all density matrices $\rho \in S_{\vec{\mu}}$. It turns out this implies that $\mathcal{L}$ is k-positive where k is the Schmidt rank of the vector $\vec{\mu}$. The property of $\vec{\mu}$-positivity implies that for all $|\psi\rangle \in S_{\vec{\mu}}$ and for arbitrary $|\phi\rangle$ we have

$$\langle\phi|(\mathbf{1} \otimes \mathcal{L})(|\psi\rangle\langle\psi|)|\phi\rangle \geq 0. \tag{A1}$$

We can always write $|\phi\rangle = (A \otimes \mathbf{1})|\Phi\rangle$ where $|\Phi\rangle$ is some maximally entangled state. By commuting A through, we obtain

$$\langle\Phi|(\mathbf{1} \otimes \mathcal{L})(|\psi_A\rangle\langle\psi_A|)|\Phi\rangle \geq 0, \tag{A2}$$

where $|\psi_A\rangle = (A^\dagger \otimes \mathbf{1})|\psi\rangle$. There is always an operator A and therefore a state $|\phi\rangle$ and $|\Phi\rangle$ such that $|\psi_A\rangle$ is proportional to the maximally entangled state with Schmidt rank k, where k is the Schmidt rank of $|\psi\rangle$ itself. Then the condition in Eq. (A2) becomes the condition for k-positivity of $\mathcal{L}$ as given in Ref. [9]. Thus $\vec{\mu}$-positivity implies k-positivity, whereas $S_{\vec{\mu}} \subset S_k$.

[1] C.H. Bennett, G. Brassard, C. Crépeau, R. Jozsa, A. Peres, and W.K. Wootters, Teleporting an unknown quantum state via dual classical and EPR channels, *Physical Review Letters* 70:1895–1899 (1993).

[2] H.-K. Lo, Classical Communication Cost in Distributed Quantum Information Processing - A generalization of Quantum Communication Complexity , quant-ph/9912009.

[3] C.H. Bennett, D.P. DiVincenzo, J.A. Smolin, B.M. Terhal and W.K. Wootters, Remote State Preparation, quant-ph/0006044.

[4] A.W. Marshall and I. Olkin, *Inequalities: Theory of Majorization and Its Applications*, Academic Press (1979).

[5] M.A. Nielsen, Conditions for a class of entanglement transformations, *Physical Review Letters*, 83:436–439 (1999), quant-ph/9811053.

[6] G. Vidal, Entanglement of pure states for a single copy, *Physical Review Letters*, 83:1046–1049 (1999), quant-ph/9902033.

[7] G. Vidal, D. Jonathan, and M. Nielsen, Approximate transformations and robust manipulation of bipartite pure state entanglement, *Physical Review A* 62:012304 (2000), quant-ph/9910099.

[8] C.H. Bennett, D.P. DiVincenzo, J.A. Smolin and W.K. Wootters, Mixed state entanglement and quantum error correction, *Physical Review A* 54:3824–3851 (1996), quant-ph/9604024.

[9] B.M. Terhal and P. Horodecki, Schmidt number for density matrices, Phys. Rev. A Rapid Communications 61:040301 (2000), quant-ph/9911117.

[10] D. Jonathan and M. Plenio, Minimal conditions for local pure-state entanglement manipulation, *Physical Review Letters* 83:1455 (1999). quant-ph/9903054

[11] A. Uhlmann, *Reports in Mathematical Physics* **45**, 407 (2000).

[12] A. Uhlmann, Simultaneous decompositions of two states, quant-ph/0007018.

[13] W.K. Wootters, Entanglement of formation of an arbitrary state of two qubits, *Physical Review Letters* 80:2245 (1998).

[14] G. Vidal, Entanglement monotones, *Journal of Modern Optics* 47:355 (2000), quant-ph/9807077.

[15] G. Vidal, Optimal local preparation of an arbitrary mixed state of two qubits; closed expression for the single copy case, quant-ph/0003002.

[16] D. Jonathan and M. Plenio, Entanglement-assisted local manipulation of pure quantum states, *Physical Review Letters* **83**, 3566–3569 (1999).